Journal of Structural Biology 185 (2014) 147–162

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Designing protein function

Modular peptide binding: From a comparison of natural binders to designed armadillo repeat proteins

Christian Reichen, Simon Hansen, Andreas Plückthun *

Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

ARTICLE INFO

Article history:
Available online 3 August 2013

Keywords:
Protein engineering
Libraries
Proteomics
Protein design
Repeat proteins
Peptide binding
Protein–peptide interactions

ABSTRACT

Several binding scaffolds that are not based on immunoglobulins have been designed as alternatives to traditional monoclonal antibodies. Many of them have been developed to bind to folded proteins, yet cellular networks for signaling and protein trafficking often depend on binding to unfolded regions of proteins. This type of binding can thus be well described as a peptide–protein interaction. In this review, we compare different peptide-binding scaffolds, highlighting that armadillo repeat proteins (ArmRP) offer an attractive modular system, as they bind a stretch of extended peptide in a repeat-wise manner. Instead of generating each new binding molecule by an independent selection, preselected repeats – each complementary to a piece of the target peptide – could be designed and assembled on demand into a new protein, which then binds the prescribed complete peptide. Stacked armadillo repeats (ArmR), each typically consisting of 42 amino acids arranged in three α-helices, build an elongated superhelical structure which enables binding of peptides in extended conformation. A consensus-based design approach, complemented with molecular dynamics simulations and rational engineering, resulted in well-expressed monomeric proteins with high stability. Peptide binders were selected and several structures were determined, forming the basis for the future development of modular peptide-binding scaffolds.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

In recent years the need for specific protein-binding molecules for experimental biology and medicine has been growing, especially for proteomic approaches. The ultimate goal is to map – ideally – every epitope, including posttranslational modifications, with specific protein detection reagents (Stoevesandt and Taussig, 2012). In sharp contrast to the increasing demand, the speed of generating such binders has not increased proportionally. Monoclonal antibodies obtained by traditional immunization are still the most frequently used binders, and their generation has essentially remained unchanged over the last 40 years. Within the last 1–2 decades, recombinant methods have provided greater control over the selection process (Plückthun et al., 2000) and revolutionized the generation of *therapeutic* proteins, notably therapeutic antibodies; in contrast, the impact on *proteomic* reagents has been modest so far. One reason is that the generation of specific binders, first in the form of antibody libraries (Knappik et al., 2000), later with alternative scaffold libraries (Binz et al., 2005), depends on the selection from a library, which has to be carried out for every target individually, and the performance and specificity of each selected binder must be evaluated individually and in detail – just as with traditional antibodies from immunization. Thus, no direct advantage is derived from having carried out previous selections: every target is a new challenge.

A modular binder, where each unit contributes to the interaction with the target molecule in a predefined manner, could overcome this limitation. Modular binding is difficult to achieve for folded protein targets. Therefore, such a modular approach is currently only conceivable for linear targets like RNA, DNA or peptides in extended conformation, and perhaps for oligosaccharides. Nucleic acids are the prime example of the use of this principle in biology: it is the prerequisite for forming a double strand. Our focus here, however, is on protein-based binding of peptides, which is a much greater challenge.

Nature has developed proteins with a modular binding mode for nucleic acids, and some of these have been technologically exploited, such as the transcription activator-like effector (TALE) repeats (Deng et al., 2012; Mak et al., 2012) or zinc-finger proteins (Klug, 2010). In both cases, repeats or domains can be linked in tandem to recognize sequence motifs in nucleic acids of various lengths. Whereas one zinc-finger domain binds three nucleotides of the DNA or RNA target, one TALE repeat recognizes one nucleotide on one strand of a dsDNA. Using this modular binding principle, artificial TALE repeat proteins were engineered to bind any consecutive nucleotide sequence of choice (Boch et al., 2009).

* Corresponding author. Address: Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. Fax: +41 44 635 57 12.
E-mail address: plueckthun@bioc.uzh.ch (A. Plückthun).

1047-8477/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsb.2013.07.012

The key question is now, how to achieve a conceptually similar modular system for peptides in extended conformation.

In the first part of this review the peptide-binding scaffolds that have been structurally described are compared and analyzed for their ability to potentially bind peptide targets in a modular fashion. To anticipate our conclusions below: we found that repeat proteins are of particular interest, since their tandem structure allows rigid stacking of repeat units and the formation of an extended structure. Among them, armadillo repeat proteins (ArmRP) form a continuous peptide-binding surface and each module binds, in a first approximation, two consecutive amino acids. An asparagine residue, conserved in almost every repeat, keeps the peptide in extended conformation through binding to the peptide backbone. Other amino acids of the binding surface provide the specificity for the target peptide by interacting with the peptide side chains. Based on these considerations, Armadillo repeat proteins (ArmRPs) were chosen as template to generate a modular peptide-binding scaffold.

In the second part of this review we describe the different protein engineering steps to develop designed ArmRPs with regular tandem repeats and favorable biophysical properties. Although modular binding along the extended binding surface of designed ArmRPs has not yet been shown, designed ArmRP have demonstrated their applicability in pull–down experiments or as detection agents in Western blots (Varadamsetty et al., 2012) and form the basis for the future development towards a modular peptide-binding system.

2. Peptide-binding: strategies and scaffolds

Protein–protein interactions are essential for living cells, and in many of these interactions both proteins are globular domains. Many of these have not only been well characterized, but also numerous protein–protein interaction scaffolds were explored for the generation of designed binding molecules as promising alternatives to traditional monoclonal or recombinant antibodies. Most of these scaffolds, such as, e.g., (in alphabetical order) Adnectins, Affibodies, Anticalins or Designed Ankyrin Repeat Proteins (DARPins) bind usually to the surfaces of folded proteins and thus do not form the focus of our review, as they have been reviewed elsewhere (Binz et al., 2005; Boersma and Plückthun, 2011; Caravella and Lugovskoy, 2010; Hosse et al., 2006; Löfblom et al., 2011; Mintz and Crea, 2013).

In the past decade, it became clear, however, that about 15–40% of all interactions in the cell are peptide–protein interactions (Petsalaki et al., 2009). By this term, we do not mean short oligopeptides, but rather unstructured regions of proteins which can bind as linear peptide targets. They can be described as unfolded terminal regions of protein domains, unstructured loops within a domain, disordered linkers between two domains (London et al., 2010) or parts of intrinsically unstructured proteins. Such transient and in general low-affinity but highly specific interactions between a globular protein and short linear peptide regions have been found in many highly dynamic cellular networks involved in signaling, regulation and protein trafficking (Diella et al., 2008; Pawson and Nash, 2003).

The challenge of binding to peptides – in contrast to stable folded proteins domains – is that peptides are usually flexible (as we are excluding structured peptides here). They thus lose a large amount of configurational entropy upon association (Killian et al., 2009). The analysis of peptide–protein binding strategies by London and coworkers (London et al., 2010) revealed that the loss of configurational entropy upon binding is minimized by the rigidity of the protein interface and compensated by an over-representation of hydrogen bonds between peptide and protein. These additional hydrogen bonds are accomplished mainly by interaction of the peptide backbone with the binding surface.

Many peptide–protein interaction domains have been characterized (Pawson and Scott, 1997; Pawson and Nash, 2003). For better comparison, we describe below all well-studied peptide-binding scaffolds according to (i) fold, (ii) target, (iii) binding mode, (iv) typical affinity and (v) natural function and illustrate them in Fig. 1. This comprehensive survey will then allow us to justify potential choices of scaffolds for the engineering of modular binding.

2.1. Antibodies

The antigen-binding variable domains (V_H and V_L) of antibodies are composed of a conserved two-β-sheet framework and six hypervariable loops, known as complementarity determining regions (CDRs) (Fig. 1A) (Sundberg, 2009). Variable in length and sequence, CDRs determine the shape of the binding site (Collis et al., 2003; MacCallum et al., 1996). Antibodies can bind to folded proteins, peptides, DNA, carbohydrates and other substances. Anti-peptide antibodies have a binding site which is usually an intermediate between the generally very deep binding pocket of anti-hapten antibodies and the relatively large and flat binding surface of protein binders (Almagro, 2004). Peptides are often bound in a groove along the dimer interface formed by the V_H and V_L domains, and sometimes one amino acid of the peptide binds in a central cavity, like a hapten. Nevertheless, the binding mode of anti-peptide antibodies is not conserved, and different peptides assume many different orientations and conformations, such as extended chains, β-turns or α-helices (Stanfield and Wilson, 1995; Sundberg, 2009). Anti-peptide antibodies can bind their target with very high affinities with reported dissociation constants in the nM range or even below (Ferrières et al., 2000; Pope et al., 2009; Luginbühl et al., 2006; Zahnd et al., 2004) and are a central component of the adaptive immune system of higher vertebrates.

2.2. MHC-I and MHC-II

The membrane-anchored and heterodimeric major histocompatibility complexes (MHC-I and MHC-II) are both composed of three domains, one peptide-binding domain and two immunoglobulin-like domains. The peptide-binding domain is composed of an eight-stranded β-sheet platform laterally enclosed by two α-helices (Yaneva et al., 2010). MHC-I proteins bind short extended peptides (8–9 amino acids in length) that originate from the intracellular degradation of (endogenous) proteins. MHC-II proteins target longer peptides (up to 20 residues) (Fig. 1A) that originate from proteolysis of engulfed extracellular (exogenous) proteins. The peptide groove of MHC-I complexes is encoded by a single protein and is closed, explaining the strict length limits of the peptide antigen (Fig. 1A). In class II MHC proteins, the binding groove is formed by two protein chains and is open at both ends, allowing MHC-II to bind longer peptides. Although MHC-II are able to bind long and highly variable peptides, the low stability and yield of MHC-II complexes makes working with them a demanding task and puts some limit on their utility as a biotechnological tool. The affinities for both complexes span a wide range. Dissociation constants between low nM and high μM have been measured (Christinck et al., 1991; Fahnestock et al., 1994; Froloff et al., 1997; Morgan et al., 1997; Sadegh-Nasseri et al., 1994). Both complexes are involved in the mammalian cellular immune response: they are presenting bound peptides at the cell surface to engage the T-cell receptor and thus activate T-cells (Neefjes et al., 2011; Rudolph et al., 2006; Vyas et al., 2008).

Download English Version:

https://daneshyari.com/en/article/2828499

Download Persian Version:

https://daneshyari.com/article/2828499

Daneshyari.com